

## Theoretical constraints on true polar wander

Victor C. Tsai<sup>1</sup> and David J. Stevenson<sup>2</sup>

Received 2 July 2005; revised 30 January 2007; accepted 22 February 2007; published 23 May 2007.

[1] For the present geologic epoch, true polar wander (TPW) is relatively small, but simple theoretical considerations suggest that it could have been larger in other epochs. In this work, we use scaling arguments to assess the qualitative behavior of TPW and a simple Maxwell model to analytically describe how changes in mass anomalies translate into TPW. Unlike previous work, we derive simple analytical estimates of TPW based on the characteristic amplitudes and timescales for changes in the moment of inertia.

We find estimates for both the amplitude and speed of TPW as a function of Earth properties. The following four main factors influence how large the maximum TPW can be: the (geological) timescale over which the desired TPW occurs  $\tau_{\text{TPW}}$ , the viscosity structure of the mantle which yields a weighted average viscosity  $\eta$ , the characteristic amplitude of the nonhydrostatic changes in the moment of inertia  $\delta C$ , and the added moment of inertia due to the equatorial bulge  $(C - A)$ . For the nominal values  $\eta = 3 \times 10^{22}$  Pa s and  $\delta C/(C - A) = 0.003$ , the maximum TPW is  $61^\circ$  over 100 Myr and  $8^\circ$  over 10 Myr. The maximum TPW speed is only sensitive to  $\eta$ ,  $\delta C$ , and  $(C - A)$ , and is  $2.4^\circ \text{ Myr}^{-1}$  for the nominal values. TPW is shown to act as a low-pass filter; rapid changes in moment of inertia produce smaller and delayed TPW. A consequence of this is that inertial interchange TPW does not have a different character than TPW. TPW can have an important contribution to plate motions over relatively long timescales but not over shorter timescales. Our simple approach allows us to assess whether multiple TPW events are possible but the major uncertainty continues to be the mantle viscosity structure.

**Citation:** Tsai, V. C., and D. J. Stevenson (2007), Theoretical constraints on true polar wander, *J. Geophys. Res.*, 112, B05415, doi:10.1029/2005JB003923.

### 1. Introduction

[2] At timescales longer than 1000 yrs, most of the Earth responds like a viscous fluid. If we keep track of the geographic location (for example, geographic latitude) of a point on the solid surface, then the following two processes can lead to large excursions of this point relative to Earth's mean angular momentum vector: Plate tectonics (i.e., mantle convection) and true polar wander (TPW). Mantle convection has received extensive study and benefits from complementary approaches that range from fundamental theory to simple scaling analyses and to detailed numerical simulations. TPW has received less attention, despite the potential to be as important in some aspects of Earth history. The basic theory is well established [Munk and MacDonald, 1960], and detailed numerical calculations exist [e.g., Ricard *et al.*, 1993; Steinberger and O'Connell, 1997; Richards *et al.*, 1999]. However, there is little analysis of an intermediate kind, firmly based in theory but devoted to scaling arguments and constraints, except at

short timescales [e.g., Johnston and Lambeck, 1999; Nakada, 2002] or at very long (fluid limit) timescales [Matsuyama *et al.*, 2006]. This paper is intended to at least partially fill this gap. It is the TPW equivalent of simple parameterized convection calculations. It is intended to guide both the users (geologists who would seek to use TPW to explain some observations and need to know what is reasonable) and the geodynamicists (who may be seeking additional theoretical constraints or intuitive understanding). We begin with a pedagogical discussion since we believe there is no adequate background of the desired kind in the current literature, either in papers or textbooks.

[3] In mantle convection, strains of order unity occur within the viscous mantle on timescales of order 100 Myr. The dissipation is comparable to the Earth's geothermal heat flow, and the differential motion of plates arises from a heat engine. By contrast, TPW involves a quasi-rigid rotation of the Earth in which the strain is  $\sim 0.003$  (Earth's oblateness), all that is required to accommodate the reorientation of the rotational bulge. The dissipation is quadratic in this strain and some five orders of magnitude or more smaller for a similar timescale and a roughly Newtonian viscous Earth. In this narrow sense, TPW is much easier than plate tectonics because the required energy budget is tiny. However, the available energy budget is also tiny. The TPW dissipation is made possible by the reduction in the Earth's rotational energy that accompanies the reorientation. The largest

<sup>1</sup>Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA.

<sup>2</sup>Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California, USA.

motions of TPW arise from processes that are ultimately the same as those driving plates; mantle convection does work building geoid anomalies, and it takes more work to build a geoid anomaly near the rotation axis than near the equator because of the centrifugal effect (other things assumed equal). This “extra” work is then released as the dissipation resulting from the reorientation to the lower rotational energy state where Earth’s rotation axis corresponds to the axis for the largest principal moment of inertia component. Events like this should arise naturally because mantle convection does not care much about Earth rotation: The Coriolis force is negligible and the degeneracy breaking arising from the centrifugal force is small. TPW also takes place because of smaller, shorter timescale processes such as ice loading and unloading. In this case, the “extra” work is provided by solar energy. In principle, TPW could also arise from other geologic processes that have no direct connection to plate tectonics and mantle convection. TPW is helped on Earth by the fragmentation of the lithosphere, which allows great circles (actually ellipses) to contract or expand in circumference by a factor  $\sim 0.003$  as the Earth reorients. This may not be so easy on current Mars, for example.

[4] A remarkable feature of TPW is not that it exists but that it is smaller for Earth than one might have supposed likely. The following argument due to the work of *Goldreich and Toomre* [1969], which still deserves attention despite all we have learned since, makes it easy to imagine that TPW motions could be considerably larger than plate tectonic motions. If the viscous relaxation time for the Earth’s rotational bulge is sufficiently short, then the Earth maintains near coincidence of the rotational axis and the axis of largest principal moment of inertia component for the part involving only convective geoid anomalies (excluding the rotational bulge). Mantle convection is chaotic at large Rayleigh number and consists of multiple ( $N$ ) cells, each of which contributes an amount  $\delta C$  incoherently to the moment of inertia tensor, and fluctuates on a convective timescale  $\tau_{\text{conv}}$ . The total anomaly will be of order  $\sqrt{N}\delta C$ , and principal axes will execute a random walk with an amplitude of order  $2\pi$  radians in time  $\sim \tau_{\text{conv}}$ . The plate tectonic motion would perhaps be  $\sqrt{N}$  smaller because each plate covers only  $1/N$  of Earth’s surface and, accordingly, has a typical size  $\sim 1/\sqrt{N}$  of Earth’s circumference. At least in recent geological time, this simple picture is not supported by observations indicating that plate motions are faster than TPW [Gordon, 1987]. There are presumably two reasons for this; (1) the relaxation is not fast enough, and (2) the motions of plates and mantle are actually more coherent than the picture described above. In all likelihood, both factors are relevant [Richards et al., 1997, 1999].

[5] TPW takes two forms. In the example given above, it is a steady reorientation of the principal axes of the residual moment of inertia tensor relative to geographic axes tied to the body of Earth. However, it is also possible (and seemingly unavoidable) that, on occasion, the difference between two of the principal moments will change sign. If  $C$ ,  $B$ , and  $A$  represent these principal moments in descending order at some epoch and if  $C - B$  changes sign at some point in time, then it will become energetically favorable for Earth to reorient through  $90^\circ$ . Such inertial interchange events, here denoted as IITPW, have been hypothesized as

major events in Earth history [Kirschvink et al., 1997; Evans, 1998; Piper, 2006; Maloof et al., 2006].

[6] In this work, for simplicity, we consider an Earth that has no secular despinning from tides and no net cooling (which changes the radius and hence average total moment of inertia) so that TPW is the result of moment of inertia temporal fluctuations that have zero mean over a long time. It is therefore possible and indeed appropriate to view this problem from the Fourier point of view and ask the following question: What is the relationship between the Fourier spectrum of the moment of inertia fluctuations and the spectrum of resulting reorientations? This is a different viewpoint than that adopted in the classic discussion by *Munk and MacDonald* [1960], where the emphasis is on the response to an isolated net change in the moment of inertia tensor. We believe that our viewpoint is not merely complementary but closer to realistic.

[7] Although there is a long and quite detailed assessment of TPW in the literature, it is not possible for an Earth scientist to take existing results and deduce readily the significance or plausibility of some hypothesized TPW event. What is the TPW “speed limit”? Could there have been much larger TPW during other geological epochs, and if so, why? Can rapid fluctuations in Earth’s moment of inertia tensor cause large TPW excursions? Is there a major difference between inertial interchange events (IITPW) and “regular” TPW? In the next section, we develop some heuristic arguments that guide us to answers of these questions. In section 3, we demonstrate the validity of these estimates and quantify them more precisely, applying the standard theory of TPW. In section 4, we discuss some applications and implications of these results.

## 2. A Heuristic Argument

[8] It is often useful to develop an understanding of a dynamical process through energy considerations, and TPW is no exception. One must also analyze the dynamical equations to confirm the heuristic picture; we do this in section 3.

[9] Consider a planet reorientation through a large angle that results in a positive degree 2 geoid anomaly being carried to the equator and the moment of inertia about the rotation axis increasing by an amount  $\delta C$ . We will refer to  $\delta C$  as a residual moment of inertia (either as a typical component amplitude or as the tensor), in analogy with the standard geodynamic meaning assigned to residual geoid. It is the part of the moment of inertia that comes from dynamic processes such as mantle convection. The angular momentum  $J$  is conserved, so the change in rotational energy is

$$\Delta E_{\text{rot}} = \frac{J^2}{2(C + \delta C)} - \frac{J^2}{2C} \approx -\frac{J^2}{2C} \cdot \frac{\delta C}{C} = -\frac{C\Omega^2}{2} \cdot \frac{\delta C}{C} \quad (1)$$

where  $\Omega$  is the angular rotation rate.

[10] Since we are dealing with a dissipative system, this reduction of rotational energy is an energy supply that is balanced by the viscous relaxation associated with reorientation of the rotational bulge (as viewed from a planet-fixed coordinate system). The strain associated with this is  $\sim \epsilon \sim \Omega^2 R^3 / GM_E \sim 0.003$ , Earth’s oblateness (where  $M_E$  is Earth’s

mass and  $R$  is Earth's radius). If the reorientation takes a time  $\tau_{\text{TPW}}$ , then the viscous energy dissipation is  $\sim \eta(e/\tau_{\text{TPW}})^2$  per unit time and volume, where  $\eta$  is the mantle viscosity. Balancing the reduction of rotational energy with the total dissipation, we get

$$\frac{1}{2}C\Omega^2 \cdot \frac{\delta C}{C} \sim \eta \left( \frac{e}{\tau_{\text{TPW}}} \right)^2 \tau_{\text{TPW}} \cdot R^3. \quad (2)$$

[11] It is useful to introduce a natural viscous relaxation time  $\tau_R = \eta/\rho g R$ , where  $\rho$  is the mantle density and  $g = GM_E/R^2$  is the gravitational acceleration. To order of magnitude, this is the timescale associated with postglacial rebound at the longest degree 2 wavelength. It then follows that

$$\frac{\tau_{\text{TPW}}}{\tau_R} \sim \frac{e \cdot C}{\delta C}. \quad (3)$$

[12] Of course, the dimensionless parameter that converts this from a scaling relationship to an actual equation does not merely involve factors like  $\pi$  but also depends on the planet structure (size of the core, variation of viscosity with depth, etc.). However, this relationship has a very simple and appealing physical interpretation. The characteristic time for large amplitude TPW is larger than the natural viscous relaxation time of the system by the ratio of the rotational bulge to the geoid anomaly driving the TPW. The planet cannot reorient through an angle larger than  $\sim \delta C/(eC)$  in a single relaxation time because the unrelaxed bulge at the previous orientation would then create a bigger (and opposing) anomaly to the moment of inertia than the driving anomaly  $\delta C$ .

[13] In practice, the two dimensionless numbers  $\tau_{\text{TPW}}/\tau_R$  and  $eC/\delta C$  are both substantially larger than unity for interesting cases. Their reciprocals provide convenient small parameters for developing a more precise theory. Any geoid anomaly at degree 2, no matter how small, could drive large TPW, as in the famous cartoon of a beetle on Earth's surface [Gold, 1955], but the timescale becomes extraordinarily long and hence impractical for small anomalies. The smallness of  $\tau_R$  also promotes isostasy and the resulting smallness of  $\delta C$ . Less obviously, this relationship also penalizes (i.e., reduces the TPW resulting from) rapid and frequent changes in the moment of inertia. In a real world, the moment of inertia is a fluctuating quantity, and if the characteristic time of the fluctuations is shorter than the time for reorientation, then that reorientation will be diminished. In this sense, TPW is a low-pass filter, allowing slow moment of inertia variations to be fully expressed in reorientations and filtering out fast variations. Unlike simple electrical circuit analogs, this first-order filtering is amplitude dependent. The Earth will continue chasing (and never quite reaching) a moving target. The “target” is the state in which the axis for the largest principal component of the residual moment of inertia tensor corresponds to the rotation axis. (This should not be confused with the question addressed and resolved by Goldreich and Toomre [1969] concerning the extent to which the nonhydrostatic part of the equatorial bulge could be attributed to an unrelaxed portion of the rotational bulge from an earlier epoch of more rapid rotation. There is no

doubt that the nonhydrostatic part of the bulge is dominated by mantle convection.)

[14] This heuristic analysis also suggests that there is no major difference in the rate of TPW between inertial interchange events (IITPW) and regular TPW. The reason is that, in both cases, the “driver” (the moment of inertia anomaly) is similar and, in both cases, the “inhibitor” (relaxation of the rotational bulge) is the same. The total change in orientation in IITPW is potentially larger. However, the rate at which it heads for that new state is no faster and can even be slower than in regular TPW, especially in the period soon after the crossover of principal components when the driver for the TPW is small. See section 4 for further analysis.

[15] The heuristic result also motivates the following obvious question: How does the characteristic timescale of TPW (the “adjustment” timescale) compare with the most important timescale that governs changes in the moment of inertia tensor (the “forcing” timescale)? It seems reasonable to think of this latter timescale as being the timescale for mantle overturn arising from thermal convection, since very high Rayleigh number convection is chaotic in character with the pattern of the convection changing almost as quickly as the overturn time. According to the simple thermal boundary layer picture, this timescale  $\tau_{\text{conv}} \sim \delta^2/\kappa$  where  $\delta$  is the thermal boundary layer thickness and  $\kappa$  is the thermal diffusivity. In accordance with the standard mantle convection scaling,  $\delta \sim 5 (\eta_{\text{conv}} \kappa / \rho g \alpha \Delta T)^{1/3}$  where  $\alpha$  is the coefficient of thermal convection and  $\Delta T$  is the temperature anomaly driving the convection. We have subscripted the viscosity to acknowledge that the averaging that enters into convective scaling is not necessarily the same averaging that enters into the viscous relaxation time introduced earlier. This formula predicts a convective timescale of 100 Ma for  $\eta_{\text{conv}} \sim 10^{21}$  Pa s,  $\kappa \sim 10^{-6}$  m<sup>2</sup>/s,  $\rho \sim 4000$  kg/m<sup>3</sup>,  $g \sim 10$  m/s<sup>2</sup>,  $\alpha \sim 2 \times 10^{-5}$  K<sup>-1</sup> and  $\Delta T \sim 500$  K. We can accordingly rewrite  $\tau_{\text{conv}} \sim 10^2 \eta_{\text{conv}} / \rho g \alpha \Delta T \delta$ . However, we also expect that the moment of inertia anomalies  $\delta C$  scale as the density anomalies  $\alpha \Delta T$ , reduced by the ratio  $\delta/R$  since these anomalies are spatially localized rather than being explicitly degree 2. Roughly speaking, this means that  $(\delta C/C) \sim 0.1 \cdot \alpha \cdot \Delta T \cdot \delta/R$ . The numerical coefficient is motivated by the tendency for anomalies to cancel and for the tendency of the direct density contribution to the moment of inertia to be canceled by the effects of dynamic topography, both at Earth's surface and at the core-mantle boundary. This is a crude estimate but it agrees roughly with the observed moment of inertia anomalies of current Earth. The important point here is not the numerical accuracy but the correct scaling with the physical parameters. Putting all these results together, we get the remarkable result

$$\frac{\tau_{\text{conv}}}{\tau_{\text{TPW}}} \sim \frac{10^2 (\eta_{\text{conv}} / \rho g \alpha \Delta T \delta)}{(10 \eta / \rho g R) \cdot (10 e R / \delta \alpha \Delta T)} \sim \frac{\eta_{\text{conv}}}{\eta e} \quad (4)$$

where a numerical factor ( $19/2 \sim 10$ ) is also inserted in the viscous relaxation time estimate (see next section). This result says that if the same viscosity estimate applies equally to convection and to viscous relaxation then convective timescales are a hundred or so times longer than TPW response timescales. It suggests that viscosity structure rather than viscosity absolute value is the central issue in



determining the vigor of TPW relative to convection. For example, if one considered an Earth with the same rotation but everywhere a lower viscosity, then both convection and TPW would be faster, but the ratio would be preserved. In reality, changes in the moment of inertia (and in mantle convection) seem to have associated timescales that are consistent with relatively low viscosity estimates of the kind that are often used in parameterized convection calculations. However, there is a considerable evidence to support rather high effective viscosities for the deep mantle [e.g., *Mitrović and Forte, 2004*], and these will play a major role in determining the effective viscosity for TPW. As a result, the difference between convective and TPW timescales may be reduced to perhaps an order of magnitude.

[16] To the extent that this kind of simplistic reasoning has merit, it motivates the view that one should not think of the forcing timescale for TPW and the adjustment timescale for TPW as being independent of each other and potentially wildly disparate, though the former is typically somewhat longer than the latter. It is consistent with the view that we have adopted that we should not think of TPW as arising from “isolated” events in Earth history but as the response to the slowly varying consequences of mantle convection that have zero mean.

### 3. Analysis

#### 3.1. Background

[17] Our analysis is an extension of the simple theory presented by *Munk and MacDonald* [1960]. The first simplifying assumption is that of a Maxwell Earth with a simple relaxation time. However, this is readily generalized to a more realistic model with multiple relaxation times, provided we are concerned with processes that operate on longer timescales than any of these relaxation times. This is indeed a good approximation for moment of inertia anomalies that arise from convection, even when they involve “rapid” plate tectonic events (for example, a slab breaking off or a reorganization of plate motions). We will identify the point in the analysis where this multiple relaxation time generalization can be applied. The absence of a lithosphere in our model is reasonable because of the fact that plate boundaries can accommodate changes in great circle lengths and that elastic effects will cancel over the timescales of interest, which are much longer than both the nutation timescale and the timescale for relative angular momentum fluctuations. Using this long timescale approximation on a Maxwell Earth and ignoring products of small quantities yields the following expression [*Munk and MacDonald, 1960*]:

$$\begin{aligned} & \frac{1}{2Q} \cdot \frac{\dot{m}_1}{m_2 m_3} + \frac{\dot{m}_2}{m_2} - \frac{\dot{m}_3}{m_3} \\ & + \theta \left( c_{33} - c_{22} + c_{13} \frac{m_1}{m_3} - c_{12} \frac{m_1}{m_2} + c_{23} \left[ \frac{m_2}{m_3} - \frac{m_3}{m_2} \right] \right) \\ & = 0 \end{aligned} \quad (5a)$$

$$\begin{aligned} & \frac{1}{2Q} \cdot \frac{\dot{m}_2}{m_1 m_3} + \frac{\dot{m}_3}{m_3} - \frac{\dot{m}_1}{m_1} \\ & + \theta \left( c_{11} - c_{33} + c_{12} \frac{m_2}{m_1} - c_{23} \frac{m_2}{m_3} + c_{13} \left[ \frac{m_3}{m_1} - \frac{m_1}{m_3} \right] \right) \\ & = 0 \end{aligned} \quad (5b)$$

$$\begin{aligned} & \frac{1}{2Q} \cdot \frac{\dot{m}_3}{m_1 m_2} + \frac{\dot{m}_1}{m_1} - \frac{\dot{m}_2}{m_2} \\ & + \theta \left( c_{22} - c_{11} + c_{23} \frac{m_3}{m_2} - c_{13} \frac{m_3}{m_1} + c_{12} \left[ \frac{m_1}{m_2} - \frac{m_2}{m_1} \right] \right) \\ & = 0 \end{aligned} \quad (5c)$$

where

$$Q = \frac{\Omega}{2\theta \cdot C}, \quad (6a)$$

$$\theta = \frac{2}{19(C - A)\tau_R}, \quad (6b)$$

$\tau_R = \eta/(\rho g R)$  is the viscous relaxation time,  $C - A$  is the excess moment of inertia due to the equatorial bulge,  $C$  is the largest moment of inertia,  $c = \delta C$  are specified perturbations in the moment of inertia tensor,  $\vec{\omega} = \Omega \vec{m}$  is the angular velocity vector, and  $\Omega$  is the magnitude of the angular velocity vector which is taken to be constant. This assumption of conservation of angular velocity is valid over the long timescales in our analysis. It should also be noted that  $c$  includes all dynamic effects due to mantle convection.

[18] At this point we recognize that it is not necessary to confine ourselves to a single relaxation time, since we can readily replace  $\tau_R$  with a sum over relaxation times and associated weightings (related to the associated love numbers, for example, as in the work of *Ricard et al.* [1993]). We then recover the same result but with a weighted  $\tau_R$ . Numerically, higher-viscosity layers tend to be weighted more heavily implying an average relaxation time weighted toward the highest relaxation time involved. Observationally, the perturbations in moment of inertia are small compared to the principal components, i.e.,  $c_{ij}/C \approx 10^{-5} \ll 1$ .

[19] A dominant balance with  $\dot{m} \sim \theta \cdot c \cdot m$  is then consistent, with the first term being smaller than the other terms by a factor of  $Q > 10^3$  and can thus be ignored. We note that this approximation will break down when one of the  $m$ 's is small, in which case we do not ignore the  $Q$  term but instead ignore the terms where the  $m$  in question is not in the denominator, which is also a consistent balance. In this degenerate case, say  $m_3 \ll 1$ , equations (5a) and (5b) then reduce to  $m_1^2 + m_2^2 \approx 1$  ( $m_3$  remains near zero) and we are left with just equation (5c) to consider. If two of three  $m$ 's are small, we have a stationary, though not necessarily stable, solution. Since there is no additional behavior of interest resulting from the degenerate cases, the remainder of this analysis focuses on equations (5a), (5b), and (5c) with the first terms ignored. All solutions presented below can be shown to satisfy equation (5a), (5b), and (5c) to the correct degree of approximation under the conditions given.

#### 3.2. Amplitude/Timescale Trade Off for TPW

[20] The easiest and simplest case to consider is when the variations in the moment of inertia involve amplitude changes and not arbitrary changes in the components of the tensor. For this analysis we transform our coordinate system to the instantaneous principal axis system (of the  $c$  tensor,

i.e.,  $c_{ij} = 0$  for  $i$  not equal  $j$ ). Letting  $c_{ii}$  vary as a function of time and introducing

$$f_{ij}(t) = c_{ii}(t) - c_{jj}(t), \quad (7)$$

equation (5a) then becomes

$$\frac{d}{dt} \left( \log \frac{m_2}{m_3} \right) = -\theta f_{32}(t), \quad (8a)$$

and integrating then gives

$$\frac{m_2}{m_3} = \left( \frac{m_2}{m_3} \right)_0 \cdot \exp \left[ -\theta \int_0^t f_{32}(t) dt \right] \quad (9a)$$

and similar expressions for equations (5b) and (5c).

[21] Although the behavior of this expression as a function of  $f_{ij}(t)$  is quite apparent, we present a Fourier analysis. Constructing a filter that maps variations in  $c_{ij}$  of different frequencies into variations in  $m_i$  allows us not only to explicitly show the dependence on the frequency of the inertia variations but also to compare with results in later sections where the full analysis is not analytically possible.

As per standard Fourier analysis,

$$f_{ij}(t) = \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega \quad (10a)$$

and

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f_{ij}(t) e^{-i\omega t} dt. \quad (10b)$$

Substituting equation (10a) into equation (9a) yields

$$\frac{m_2}{m_3} = \left( \frac{m_2}{m_3} \right)_0 \cdot \exp \left[ \theta \cdot i \int_{-\infty}^{\infty} \frac{F(\omega)}{\omega} e^{i\omega t} d\omega \right] = u(t). \quad (11a)$$

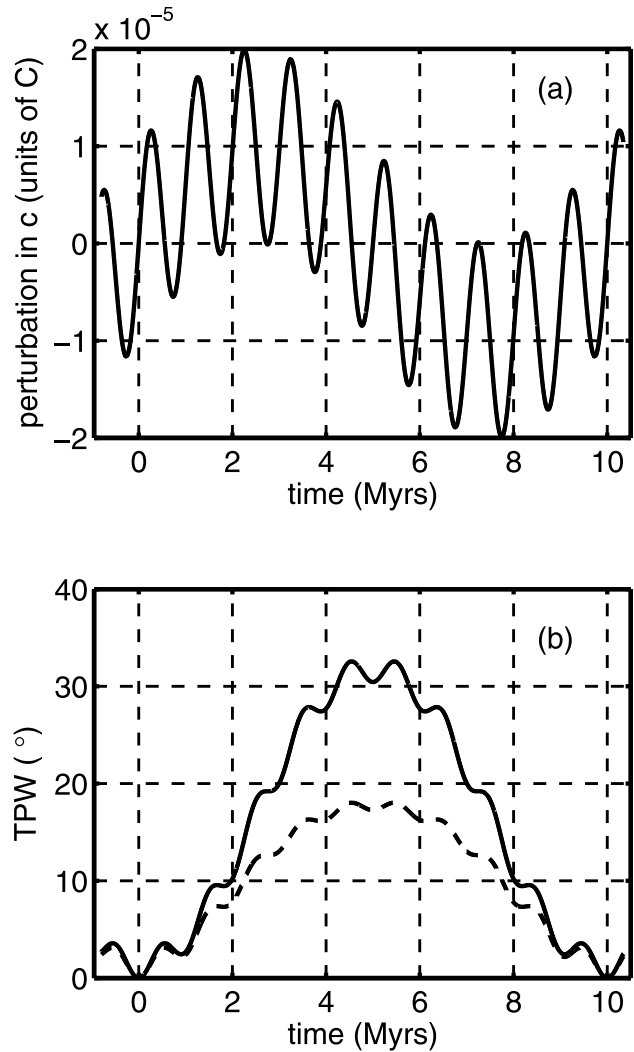
The important aspect of this result is that the dependence of  $m_2/m_3$  on  $f_{32}(t)$  is that of a  $1/\omega$  filter. That is, a variation (with a prescribed amplitude) with a period twice as long will achieve twice as large of a change in  $\log(m_2/m_3)$ . Consequences of this  $1/\omega$  filter are that short-term variations are smoothed out. For example, Figure 1a shows a variation in  $f_{32}(t)$  with two dominant frequencies, and Figure 1b shows the response (amount of TPW).

[22] To show the dependence of the maximum TPW on  $\omega$  more explicitly, we look at the response due to a single sinusoid of a given frequency. That is, we let  $f_{ij}(t) = f_{32} \cdot \cos(\omega t)$ . We also introduce  $\varphi$  so that  $\tan \varphi = m_i/m_j = u(t)$  and let  $(m_2/m_3)_0 = \exp(n)$ . Then equation (11a) simplifies to

$$\tan \varphi = \exp \left[ n + \frac{\theta f_{32}}{\omega} \sin \omega t \right]. \quad (12)$$

The maximum and minimum values for  $\varphi$  occur at  $t = \pi/2$  and  $t = -\pi/2$ . Substituting these values gives

$$\tan \Delta \varphi = \tan(\varphi_1 - \varphi_2) = \text{sech } n \cdot \sinh \left[ \frac{\theta f_{32}}{\omega} \right] \quad (13)$$



**Figure 1.** (a) Fluctuations in the moment of inertia ( $c_{ij}$ ) given by the sum of two sinusoids of different periods but similar amplitudes. (b) Two possible TPW response functions due to the forcing in Figure 1a. The solid curve is the maximum TPW curve. Note that, in both cases, the high-frequency perturbation is filtered out. (The solid curve begins at  $\varphi \sim 67^\circ$ , and the dashed curve begins at  $\varphi \sim 27^\circ$ .)

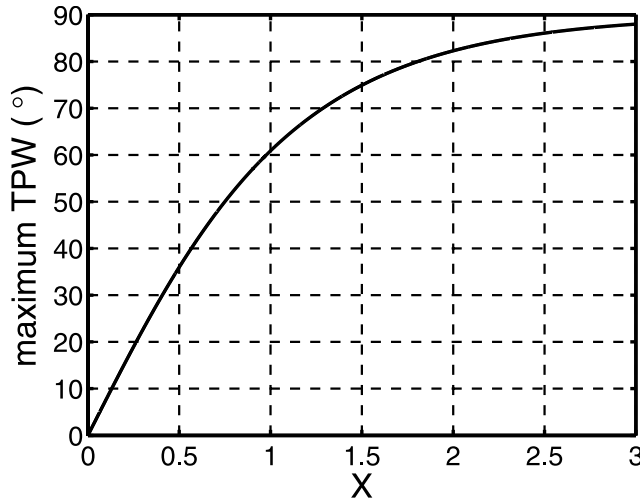
which obtains a maximum at  $n = 0$ , so the maximum TPW is given by

$$\Delta \varphi_{\max} = \tan^{-1} \left[ \sinh \left( \frac{\theta f_{32}}{\omega} \right) \right] = \tan^{-1} \left[ \sinh \left( \frac{2}{19} \cdot \frac{c_{33} - c_{22}}{C - A} \cdot \frac{1}{\omega \tau_R} \right) \right] \quad (14)$$

or with  $\omega \tau_{\text{TPW}} = 2\pi$  and introducing the dimensionless variable  $X$ , scaled so that  $X \sim 1$  when  $c_{ij} = 10^{-5}C$ ,  $(C - A)/C = 300$ ,  $\tau_{\text{TPW}} = 10^8$  yrs,  $\eta = 3 \times 10^{22}$  Pa s (roughly equivalent to  $\tau_R \sim 10^3$  yrs), then:

$$X = \frac{c_{33} - c_{22}}{0.003(C - A)} \cdot \frac{\tau_{\text{TPW}}}{100 \text{ Myr}} \cdot \frac{3 \cdot 10^{22} \text{ Pa} \cdot \text{s}}{\eta} \quad (15)$$

$$\Delta \varphi_{\max} = \tan^{-1} [\sinh(1.3X)] \quad (16)$$



**Figure 2.** The maximum TPW as a function of  $X$  where  $X = (c_{33} - c_{22})/[0.003(C - A)] \cdot \tau_{\text{TPW}}/100 \text{ Myr} \cdot 3 \times 10^{22} \text{ Pa s}/\eta$ , as described by equation (15) in section 3.2.

[23] This result is plotted in Figure 2. One should note that the expression has the same general form as the heuristic argument suggests. See section 4 for a more thorough explanation for the scaling choices.

[24] As well as a result for the maximum TPW, we also obtain a speed limit for TPW that is the maximum rate at which TPW can occur. Rewriting equation (8a) with  $\tan\phi = m_i/m_j = u$  gives

$$\dot{\phi} = -\theta f_{32}(t) \frac{u}{1 + u^2}. \quad (17)$$

We maximize this TPW speed as a function of  $u$  and obtain:

$$|\dot{\phi}_{\text{max}}| = \left| \frac{1}{2} \theta f_{32}(t) \right| = \frac{2.4^\circ}{1 \text{ Myr}} \cdot \left| \frac{c_{33}(t) - c_{22}(t)}{0.003(C - A)} \cdot \frac{3 \cdot 10^{22} \text{ Pa} \cdot \text{s}}{\eta} \right| \quad (18)$$

which is attained when  $u(t) = \pm 1$ ,  $m_i = \pm m_j$ , or  $\phi = \pm\pi/4$ . See section 4 for numerical answers.

### 3.3. Adding Off-Diagonal Terms

[25] To perform a complete analysis of the problem, we need to be able to explain how to account for arbitrary variations in the moment of inertia tensor (not just variations in the principal components in some reference frame). As the following analysis shows, the results with off-diagonal terms are not significantly different from the results in the previous section, although they do not have a clean closed-form solution.

[26] Since it is difficult to solve equations (5a), (5b), and (5c) in a manner that provides insight into the problem, we instead describe solutions to equations (5a), (5b), and (5c) without actually solving the equations. One cannot just use equation (9a) and re-diagonalize the  $\mathbf{c}$  tensor at every time step because equation (9a) is not the solution when the variations in  $\mathbf{c}$  are not aligned with the instantaneous principal directions. Instead, we allow for a nondiagonal  $\mathbf{c}$  but choose a basis such that the changes in  $\mathbf{c}$  are in principal directions. In other words, for motions over time  $\Delta t$ , we have

$$\mathbf{c}(t + \Delta t) = \mathbf{c} + \Delta \mathbf{c}. \quad (19)$$

[27] Choosing the eigenvectors of  $\Delta \mathbf{c}$  as our new basis gives us a new tensor  $\mathbf{c}$  for which the changes instantaneously only occur along principal components.

[28] Since it is still difficult to gain insight with all the nondiagonal terms present, we simplify the problem further as follows: We allow a nonzero but constant  $c_{12}$  component and varying  $c_{ii}(t)$  components. Because of the symmetry in equation (5c) (but not in equations (5a) or (5b)) from the choice of nonzero  $c_{ij}$ , we examine this equation first.

[29] Letting  $u(t) = m_2/m_1$  (and again ignoring the  $Q$  terms), equation (5c) then simplifies to

$$\frac{\dot{u}}{u} + \theta \left( f_{12}(t) + c_{12} \left( u - \frac{1}{u} \right) \right) = 0. \quad (20)$$

It is easier to obtain a speed limit in this case than to obtain the full solution so we present this first. To obtain the speed limit, we again replace  $\tan\phi = u$ . But

$$u - \frac{1}{u} = -2 \cot \psi \quad (21a)$$

and

$$\frac{\dot{u}}{u} = \frac{\dot{\psi}}{\sin \psi} \quad (21b)$$

where  $\psi = 2\phi$  so that equation (20) becomes

$$\dot{\psi} = -\theta f_{12}(t) \sin \psi + 2\theta c_{12} \cos \psi. \quad (22)$$

Maximizing this TPW speed as a function of  $\psi$ , we obtain:

$$|\dot{\psi}_{\text{max}}| = \frac{|\dot{\psi}_{\text{max}}|}{2} = \frac{\theta}{2} \sqrt{f_{12}^2(t) + 4c_{12}^2} \quad (23)$$

which is attained when

$$\sin \psi = \frac{f_{12}}{\sqrt{f_{12}^2 + 4c_{12}^2}}, \quad (24a)$$

$$\cos \psi = \frac{2c_{12}}{\sqrt{f_{12}^2 + 4c_{12}^2}}. \quad (24b)$$

[30] One should note the fact that equation (23) has exactly the same form as equation (18) and is applicable with a time varying  $c_{12}$  term as well as a constant one. It should also be noted that the factor of 4 arises not because the nondiagonal terms are more important but because the inertia tensor is symmetric. For example,  $c_{12} = 1$  is equivalent to  $c_{1'1'} = 1$ ,  $c_{2'2'} = -1$  and thus  $f_{1'2'} = 2$  in a principal frame (with axes  $1' 2' 3'$ ).

[31] Continuing to differentiate equation (22), one obtains

$$\ddot{\psi} = \frac{\theta^2}{2} ([f_{12}^2(t) - 4c_{12}^2] \sin 2\psi - 4f_{12}(t) \cdot c_{12} \cos 2\psi) - \theta \cdot \dot{f}_{12}(t) \sin \psi. \quad (25)$$

If we include a time varying  $c_{12}$  term, then we obtain:

$$\ddot{\psi} = \frac{\theta^2}{2} ([f_{12}^2(t) - 4c_{12}^2] \sin 2\psi - 4f_{12}(t) \cdot c_{12} \cos 2\psi) - \theta \cdot \dot{f}_{12}(t) \sin \psi + 2\theta \cdot \dot{c}_{12} \cos \psi. \quad (26)$$

Equation (26) has the following two obvious limits: that of long and short forcing timescales. If the forcing has the characteristic amplitude  $f$  and timescale  $\tau$ , then for

$$\tau \gg \frac{1}{\theta f} \quad (27a)$$

and

$$\tau \ll \frac{1}{\theta f} \quad (27b)$$

we, respectively, have for the maximum TPW acceleration:

$$|\ddot{\phi}_{\max}| = \frac{\theta^2}{8} (f_{12}^2(t) + 4c_{12}^2) \quad (28a)$$

and

$$|\ddot{\phi}_{\max}| = \frac{\theta}{2} \sqrt{f_{12}^2(t) + 4c_{12}^2}. \quad (28b)$$

One should note that, for a forcing with multiple timescales (of similar amplitudes), the timescale in question here is the shortest one, whereas for the amplitude of TPW the longest timescale is the most important. One should also note that our analysis yields a lower bound for the maximum TPW acceleration. That is, regardless of timescale, the (theoretical, not necessarily observed) maximum TPW acceleration is always at least as large as the quantity given in equation (28a).

[32] Now we present the approximate solution for  $u(t)$ . For  $f_{12}(t) = f_{12}$  constant, we solve equation (20) analytically as follows:

$$u = \frac{1}{2c_{12}} \left( f_{12} + \sqrt{f_{12}^2 + 4c_{12}^2} \tanh \left[ \frac{\theta}{2} \sqrt{f_{12}^2 + 4c_{12}^2} (t - t_0) \right] \right). \quad (29)$$

For this constant  $f_{12}$  case, the maximum TPW amplitude should be  $\Delta\phi_{\max} = \pi/2$  since it corresponds to the infinite timescale region (that is,  $\omega$  approaches zero), and indeed, for all values of  $c_{12}$  and  $f_{12}$ ,  $\Delta\phi_{\max} = \pi/2$  and the timescale of reorientation is  $1/b$  where  $b = |\dot{\phi}_{\max}|$ . Moreover, using (hyperbolic and standard) trigonometric identities, it can be shown that the TPW follows the exact same functional form as the diagonal case. This should not be surprising since the two cases are identical up to a change of coordinates.

[33] For nonconstant  $f_{12}(t)$ , the solution is nearly identical to either the diagonal case or the constant  $f_{12}$  case, discussed in the previous paragraph. For large  $f_{12}(t)$  compared to  $c_{12}$ , the diagonal frame is nearly identical to the nondiagonal frame and thus the solution is comparable to the diagonal solution (12). If, on the other hand,  $f_{12}(t)$  is small compared to  $c_{12}$ , then the perturbation on top of the solution (29) is small, and again, we have a similar behavior. Thus the addition of the  $c_{12}$  term does not qualitatively change the solution. We can, however, do some approximate analysis that shows this behavior more precisely. For the case where  $f_{12}(t)$  is small compared to  $c_{12}$ , we can use a perturbative approximation and variation of parameters, taking only one Fourier component, and obtain to first order

$$z = \frac{1}{z_0} \left( 1 + \frac{f_{12}^2}{8c_{12}^2} \right) \tanh[b(t - t_0)] + \frac{f_{12}^2 \theta}{16c_{12}(\omega^2 + b^2)} \times (b \cos 2\omega t \cdot \tanh[b(t - t_0)] + \omega \sin 2\omega t) \quad (30)$$

where

$$z = u \frac{f_{12}}{2c_{12}} \quad (31a)$$

and

$$z_0 = \sqrt{1 + \left( \frac{f_{12}}{4c_{12}} \right)^2} = \frac{1}{2c_{12}} \sqrt{f_{12}^2 + 4c_{12}^2}. \quad (31b)$$

Comparing equation (30) with equation (29), the corrections are second order in  $x = f_{12}/(2c_{12})$ . Furthermore, these corrections are of the form of equation (12), except with  $\sqrt{\omega^2 + b^2}$  instead of  $\omega$ .

[34] Now with an approximate solution to equation (5c), we can substitute the result for  $m_2/m_1$  into equations (5a) and (5b), where this term acts as a linearly added forcing term that can therefore be treated as an addition to the  $f_{32}(t)$  and  $f_{13}(t)$  terms, respectively. Although  $m_2/m_1$  may not vary sinusoidally, it can be Fourier decomposed so that all of the analysis in section 3.2 still holds. Moreover, the dominant Fourier components of this forcing will be close to that of the original forcing  $f_{12}(t)$  except with different amplitudes (as seen from the form of the solution), with other smaller components added due to the nonsinusoidal character of the response. One can therefore treat these effects as contributing a similar forcing to the two equations where we lack symmetry. Using these solutions in an iterative sense, we can then construct the solution with all off-diagonal terms present.

[35] Thus while we began this discussion with only one nonzero off-diagonal component and only described the solution to one of the three governing equations, we now have a qualitative description of the full nonlinear system. There are some obvious caveats to this description. First, as was noted, we have only done the analysis in a perturbative approximation. We argue that, since the behavior is similar on both sides of the approximation, it should approximately hold for intermediate values as well. However, only by performing a full nonlinear analysis (which is beyond the scope of this paper) can this claim be truly substantiated. Second, the exact form of the secondary forcing as described in the above paragraph is not easily described, and therefore, the result is not as clear as in the diagonal case. Finally, it is not obvious that the iterative application of these results yields a result similar to the diagonal case, although the arguments above suggest that there is no qualitative difference.

[36] We have shown that both the TPW speed and maximum amplitudes show similar behaviors for the off-diagonal and principal component cases. Moreover, the qualitative behavior of the equations does not seem to change. We use these facts to conclude that we can apply our results for the principal component case to a general variation in the moment of inertia tensor, with the important caveat that, in order to apply the earlier results, one must still view the problem in a coordinate system such that the changes are only in the principal directions.

### 3.4. Summary of Analysis

[37] Given our conclusions about the general applicability of the diagonal case, the three most important primary results that should be remembered are equations (9a), (16) [along with equation (15)], and (18). They represent the



following: [equation (9a)] the TPW as a function of an arbitrary variation in the diagonal components of the moment of inertia tensor; [equation (16)] the maximum TPW as a function of (1) the amplitude and (2) timescale of variation in the diagonal components, (3) the average mantle viscosity, and (4) the added moment of inertia due to the hydrostatic bulge; [equation (18)] the maximum TPW velocity as a function of (1) amplitude of variation in the diagonal components, (2) the average mantle viscosity, and (3) the added moment of inertia due to the bulge. Thus, the quantities of importance are the amplitudes of  $C - A$ ,  $\eta$ , and  $c_{ij}$ . Also of importance is  $\omega$  or, more precisely, the frequency content of  $f_{32}(t) = c_{33}(t) - c_{22}(t)$ .

[38] In addition to these results, we obtained expressions for the TPW Fourier decomposition of the diagonal case, the maximum TPW velocity for the full nonlinear case (no approximation), the maximum TPW acceleration for the full nonlinear case (no approximation), and an approximate TPW for a single frequency forcing in the off-diagonal case (see sections 3.2 and 3.3).

#### 4. Realistic Parameter Choices

[39] In order to assess the expected TPW history of the Earth we need to know several parameters and how they may have varied through time: the residual moment of inertia tensor typical amplitude, the characteristic timescale (or equivalently spectral content) of this tensor, the mantle relaxation time(s), and the rotation rate.

[40] The typical amplitude of  $c_{ij}/C$  is unlikely to be significantly different in the past than it is at present, i.e.,  $10^{-5}$ . Although this value can be approximately predicted by existing mantle convection models, it is not merely the consequence of the expected density anomalies arising from thermal convection. It also depends on the dynamic topography created at the Earth's surface and at internal interfaces (for example, the core-mantle boundary and possibly other interfaces in the mantle). As is well known from the study of the Earth's geoid [Hager *et al.*, 1985], dynamic topography effects are comparable to the direct effect of density anomalies and often change the sign of the net geoid. Nonetheless, the typical total moment of inertia anomaly can still be understood to order of magnitude as arising from the typical fractional amplitude of density anomalies  $\alpha\Delta T \sim 10^{-3}$  and the fact that these anomalies are confined to regions of order a few percent (for example, the thickness of slabs divided by the depth of the mantle). Since the degree 2 geoid can be thought of as arising from many contributions, roughly as many of one sign as of the opposite sign, there must presumably be epochs where the amplitude is a factor of a few larger than now as well as epochs where particular elements of the residual tensor pass through zero.

[41] It is important to realize that individual "events" in Earth's plate tectonic history are unlikely to have effects that are larger than this mean current RMS level of  $c_{ij}$ . To appreciate this, consider the following two examples:

[42] (1) A major plume rising from the deep mantle. For example, a (very large) blob of radius  $R_p = 1000$  km and temperature anomaly  $\Delta T = 300$  K will correspond to a mass anomaly  $M \sim 4\pi\rho \cdot R_p^3 \cdot \alpha \cdot \Delta T/3 \sim 5 \cdot 10^{-6} M_E$ , assuming  $\alpha \sim 1 \times 10^{-5} \text{ K}^{-1}$  is appropriate to deep mantle. Even allowing for the potentially larger effects of boundary

deformations, this suggests a typical moment of inertia anomaly that changes by less than one part in  $10^5$  over the time of ascent.

[43] (2) A slab that breaks off could produce a change in moment of inertia comparable to its total density anomaly (because of the change in dynamic topography arising from the change in mantle flow regime). For example, a slab with a horizontal length of 5000 km, vertical extent of 1000 km, and thickness of 100 km with temperature anomaly of 600 K would have a comparable mass anomaly as in the plume example above.

[44] The biggest problem with estimates of TPW lies in the determination of mantle viscosity structure. This most directly affects the estimate of relaxation time (here lumped into a single time, or equivalently lumped into a single average viscosity as described in section 3.1). However, it also enters indirectly through affecting the mantle convection timescale and (perhaps) the long-term stability of the convection. In our heuristic discussion we showed that there may be a sense in which viscosity cancels because it affects both the relaxation timescale and the mantle convection timescale in somewhat similar ways. However, this is a potentially misleading trade off since a very high lower mantle viscosity may also reduce the level of fluctuation in the mantle convection, thus lowering the magnitude of moment of inertia fluctuations. For a nominal mantle viscosity of  $3 \times 10^{22} \text{ Pa s}$ , the effective relaxation time ( $19/2 \times \tau_R$ ) is only  $4 \times 10^4$  yrs.

[45] Earth's rotation has changed through geologic time as tides transfer from spin angular momentum to orbital angular momentum in the Moon's orbit. In the Archean, the shorter length of day means a larger rotational bulge and an accordingly stabilizing effect on TPW. It might be supposed that the higher heat flow and faster mantle convection of earlier epochs would promote more rapid TPW, and this is indeed expected. However, from our heuristic analysis, we observed that the underlying control parameter, mantle viscosity, enters in a similar way to reduce the TPW timescale and the mantle convective timescale at about the same rate. Thus the behavior of TPW at earlier epochs is roughly similar to now except for the greater rapidity. The change in rotation rate is smaller than the expected change in viscosity. For example, a rotation rate of 15 hours instead of 24 hours changes  $e$  by a factor of a little more than two. In the same period (3 billion yrs) it is reasonable to expect a change in mantle viscosity of about an order of magnitude.

[46] Finally, we comment on the implications of these realistic choices for some proposed TPW events. A number of workers, especially in the paleomagnetic community, seem to have the idea that inertial interchange TPW (IITPW) can be significantly faster than regular TPW. The idea seems to be that IITPW events occur immediately after the intermediate moment of inertia becomes larger than the largest moment of inertia and that, if the switch occurs again, another IITPW event can occur. This is not the case. For example, in Figure 1, this switch in axes occurs 10 times in 10 Myr but only one large (up to  $33^\circ$ ) TPW excursion occurs. A specific IITPW scenario that has been argued for by Kirschvink *et al.* [1997] is what we interpret to be roughly  $80^\circ$  in 40 Myr. We choose less extreme values than claimed by the authors since IITPW (TPW of  $90^\circ$ ) is only approached asymptotically (in the diagonal case). Assuming



**Table 1.** Possible Choices of Parameters and the Resulting Maximum TPW<sup>a</sup>

| $\delta C/C$ | $\tau_{\text{TPW}}$ (Myrs) | $\eta$ (Pa·s)      | $X$  | $\Delta\phi_{\text{max}}$ |
|--------------|----------------------------|--------------------|------|---------------------------|
| $10^{-5}$    | 300                        | $3 \times 10^{22}$ | 3    | $88^\circ$                |
| $10^{-5}$    | 100                        | $3 \times 10^{22}$ | 1    | $61^\circ$                |
| $10^{-5}$    | 10                         | $3 \times 10^{22}$ | 0.1  | $8^\circ$                 |
| $10^{-5}$    | 1                          | $3 \times 10^{22}$ | 0.01 | $0.8^\circ$               |
| $10^{-5}$    | 100                        | $10^{23}$          | 0.3  | $23^\circ$                |
| $10^{-6}$    | 10                         | $3 \times 10^{22}$ | 0.01 | $0.8^\circ$               |

<sup>a</sup>Note the linear dependence of  $X$  on each factor. The reader can easily substitute their favorite values of parameters to obtain an  $X$  that gives the maximum TPW as in Figure 2.  $\delta C/C$  is the scaled residual moment of inertia,  $\tau_{\text{TPW}}$  is the timescale of interest,  $\eta$  is the weighted average mantle viscosity,  $X$  is the input parameter to equation (16), and  $\Delta\phi_{\text{max}}$  is the maximum TPW for the given parameter choice.

the largest physically reasonable value for the change in moment of inertia ( $\delta C/C = 10^{-5}$ ), the scenario still requires  $\eta = 7 \times 10^{21}$  Pa s (or smaller) in order for such TPW to occur, which is unlikely to be true for the Earth. In this case, the authors should rethink their conclusions since the results are dynamically implausible. Others, like *Li et al.* [2004], also argue for IITPW but have data that point to less-drastic TPW of what we estimate to be  $70^\circ$  in 60 Myr. With the same assumptions as above, this requires  $\eta = 2 \times 10^{22}$  Pa s (or smaller), which is perhaps closer to the range of plausible values for the average lower mantle viscosity.

## 5. Conclusions

[47] Our heuristic analysis shows that the timescale for large amplitude (close to  $90^\circ$ ) TPW is larger than the natural viscous relaxation timescale by the ratio of the hydrostatic bulge to the geoid anomaly driving the TPW (see section 2). Our more complete analysis shows that there are four main factors that strongly influence how large the theoretical maximum TPW can be. All four of these factors enter either linearly or inversely into determining the factor  $X$  described in section 3.2, where  $X$  in turn completely describes the maximum TPW that arises, through equation (16) plotted in Figure 2. The amplitude and timescale of fluctuations in the moment of inertia both enter linearly, whereas the extra moment of inertia of the hydrostatic bulge ( $C - A$ ) and the (weighted) average mantle viscosity both enter inversely. Some values are given in Table 1. These results show that TPW can be faster than plate tectonic motions (which lead to apparent polar wander). For example, in only 10 Myr, we can achieve a TPW of  $8^\circ$  (with  $\eta \sim 3 \times 10^{22}$  Pa s). However, in order to achieve values of TPW as high as  $88^\circ$  in 10 Myr, an average viscosity of  $\eta \sim 10^{21}$  Pa s would be required. This is below the current minimum estimates of average viscosity (see section 4) and thus is not a reasonable amount of TPW to expect in 10 Myr. With this said, it should be noted that the average mantle viscosity is not well constrained. We showed (see section 2) that the timescale for TPW and mantle convection scale in a similar way with mantle viscosity, with TPW predicted to be somewhat faster. However, the relatively unknown viscosity structure makes it difficult to present more than a scaled estimate of the theoretical maximum TPW (and TPW rate; see section 4 for details).

[48] The maximum TPW speed is given by equation (18) and is  $2.4^\circ \text{ Myr}^{-1}$  multiplied by a factor of order 1. (This is

equivalent to  $27 \text{ cm/yr.}$ ) However, this maximum rate is only achieved for a relatively small length of time, in the middle of a TPW “event.” Thus the total TPW cannot be viewed as the maximum TPW rate multiplied by the timescale in question. Moreover, it should be stressed that actual changes in moment of inertia are seldom abrupt, and so, it is useful to view the problem as a sum of Fourier contributions rather than as a sum of impulsive contributions. Under this framework, TPW can be viewed as a low-pass filter. Rapid changes in moment of inertia produce subdued and delayed responses. The biggest drivers of TPW are long timescale changes.

[49] Another point that our analysis makes clear is that inertial interchange TPW (IITPW) of the sort that, for example, *Kirschvink et al.* [1997] and *Evans* [1998] discussed do not have a different character than conventional TPW. The same processes (i.e., mantle convection) are responsible for both so IITPW does not produce faster TPW. In addition, IITPW is delayed in time by the gradual growth of the moment of inertia driving force as described above.

[50] Finally, it is of interest that, for much of geologic time, the evidence supports less TPW than our simple model predicts as the maximum value. There are a few possible reasons for this, of which we comment on three. (1) It may be that there is an unusual or surprising coherence of mantle convection/plate tectonics. A possible reason for this that has not been previously considered is the possible stabilizing effects that the 0.3% topography difference between pole and equator may have on mantle convection. (2) The lower mantle may be very viscous. As described in section 3.1, the higher viscosity layers are weighted more. The lower mantle viscosity is not well constrained by any current observations. Thus if the lower mantle is many times more viscous (for example, 100 times) than the upper mantle, then we would have a much smaller maximum TPW angle. (3) Possible triggers for TPW may be too fast or too small. It may be that events that would cause large TPW are short-lived processes that are accordingly damped out.

[51] **Acknowledgments.** We thank G. Acton and an anonymous reviewer for helpful comments. This research was partially supported by a National Science Foundation Graduate Fellowship (VCT).

## References

- Evans, D. A. (1998), True polar wander, a supercontinental legacy, *Earth Planet. Sci. Lett.*, *157*, 1–8.
- Gold, T. (1955), Instability of the Earth’s axis of rotation, *Nature*, *175*, 526–529.
- Goldreich, P., and A. Toomre (1969), Some remarks on polar wandering, *J. Geophys. Res.*, *74*, 2555–2567.
- Gordon, R. G. (1987), Polar wandering and paleomagnetism, *Annu. Rev. Earth Planet. Sci.*, *15*, 567–593.
- Hager, B. H., R. W. Clayton, M. A. Richards, R. P. Comer, and A. M. Dziewonski (1985), Lower mantle heterogeneity, dynamic topography and the geoid, *Nature*, *313*, 541–545.
- Johnston, P., and K. Lambeck (1999), Postglacial rebound and sea level contributions to changes in the geoid and the Earth’s rotation axis, *Geophys. J. Int.*, *136*, 537–558.
- Kirschvink, J. L., R. L. Ripperdan, and D. A. Evans (1997), Evidence for a large-scale reorganization of early Cambrian continental masses by inertial interchange true polar wander, *Science*, *277*, 541–545.
- Li, Z. X., D. A. D. Evans, and S. Zhang (2004), A  $90^\circ$  spin on Rodinia: Possible causal links between the Neoproterozoic supercontinent, superplume, true polar wander and low-latitude glaciation, *Earth Planet. Sci. Lett.*, *220*, 409–421.

- Maloof, A. C., G. P. Halverson, J. L. Kirschvink, D. P. Schrag, B. P. Weiss, and P. F. Hoffman (2006), Combined paleomagnetic, isotopic, and stratigraphic evidence for true polar wander from the Neoproterozoic Akademikerbreen Group, Svalbard, Norway, *Geol. Soc. Am. Bull.*, *118*, 1099–1124.
- Matsuyama, I., J. X. Mitrovica, M. Manga, J. T. Perron, and M. A. Richards (2006), Rotational stability of dynamic planets with elastic lithospheres, *J. Geophys. Res.*, *111*, EO2003, doi:10.1029/2005JE002447.
- Mitrovica, J. X., and A. M. Forte (2004), A new inference of mantle viscosity based upon joint inversion of convection and glacial isostatic adjustment data, *Earth Planet. Sci. Lett.*, *225*, 177–189.
- Munk, W. H., and G. J. F. MacDonald (1960), *The Rotation of the Earth*, Cambridge Univ. Press, New York.
- Nakada, M. (2002), Polar wander caused by the quaternary glacial cycles and fluid love number, *Earth Planet. Sci. Lett.*, *200*, 159–166.
- Piper, J. D. A. (2006), A ~90 degree late Silurian-early Devonian apparent polar wander loop: The latest inertial interchange of planet Earth?, *Earth Planet. Sci. Lett.*, *250*, 345–357.
- Ricard, Y., G. Spada, and R. Sabadini (1993), Polar wandering of a dynamic Earth, *Geophys. J. Int.*, *113*, 284–298.
- Richards, M. A., Y. Ricard, C. Lithgow-Bertelloni, G. Spada, and R. Sabadini (1997), An explanation for Earth's long-term rotational stability, *Science*, *275*, 372–375.
- Richards, M. A., H. P. Bunge, Y. Ricard, and J. R. Baumgardner (1999), Polar wandering in mantle convection models, *Geophys. Res. Lett.*, *26*, 1777–1780.
- Steinberger, B., and R. J. O'Connell (1997), Changes of the Earth's rotation axis owing to advection of mantle density heterogeneities, *Nature*, *387*, 169–173.

---

D. J. Stevenson, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA.

V. C. Tsai, Department of Earth and Planetary Sciences, Harvard University, 20 Oxford St., Cambridge, MA 02138, USA. (vtsai@fas.harvard.edu)